

0590
0328

#7



ENTERED

OIPE

RAW SEQUENCE LISTING

DATE: 04/01/2003

PATENT APPLICATION: US/10/081,739A

TIME: 13:17:23

Input Set : A:\09010-107001.txt

Output Set: N:\CRF4\04012003\J081739A.raw

```

4 <110> APPLICANT: Callen, Walter
5      Richardson, Toby
6      Frey, Gerhard
7      Miller, Carl
8      Kazaoka, Martin
9      Short, Jay
10     Mathur, Eric
12 <120> TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY
13     AND METHODS OF USE THEREOF
15 <130> FILE REFERENCE: 09010-107001
17 <140> CURRENT APPLICATION NUMBER: 10/081,739A
18 <141> CURRENT FILING DATE: 2002-02-21
20 <150> PRIOR APPLICATION NUMBER: 60/270,495
21 <151> PRIOR FILING DATE: 2001-02-21
23 <150> PRIOR APPLICATION NUMBER: 60/270,496
24 <151> PRIOR FILING DATE: 2001-02-21
26 <150> PRIOR APPLICATION NUMBER: 60/291,122
27 <151> PRIOR FILING DATE: 2001-05-14
29 <160> NUMBER OF SEQ ID NOS: 69
31 <170> SOFTWARE: FastSEQ for Windows Version 4.0
33 <210> SEQ ID NO: 1
34 <211> LENGTH: 1311
35 <212> TYPE: DNA
36 <213> ORGANISM: Artificial Sequence
38 <220> FEATURE:
39 <223> OTHER INFORMATION: Synthetically generated
41 <400> SEQUENCE: 1
42 atggccaagt actccgagct ggaaaagggc ggggtcataa tgcaggcggt ctactgggac      60
43 gtgccttcag gaggaatatg gtgggacaca ataccggcaga agataccgga gtggtacgat      120
44 gccggaatct ccgcaatatg gattcccccg gcgagcaagg gcatgggcgg cgcctattcg      180
45 atgggctacg acccctacga cttctttgac ctcggtgagt acgaccagaa gggaacggta      240
46 gagacgcgct ttggctccaa gcaggagctc gtgaacatga taaacaccgc ccacgcctat      300
47 ggcatgaagg taatagccga tatagtcata aaccaccgcg ccggcggtga cctggagtgg      360
48 aacccttcg tgaacgacta tacctggacc gacttctcaa aggtcgcgtc gggtaaatac      420
49 acggccaact acctcgactt ccaccgaac gagctccatg cgggcgattc cggaacattt      480
50 ggaggctatc ccgacatatg ccacgacaag agctgggacc agtactggct ctgggccagc      540
51 caggagagct acgcggcata tctcaggagc atcggcatcg atgcctggcg cttcgactac      600
52 gtcaagggct atgctccctg ggtcgtcaag gactggctga actggtgggg aggctgggag      660
53 gttggagagt actgggacac caacgtcgac gctgttctca actgggcata ctcgagcggt      720
54 gccaaggtct ttgacttcgc cctctactac aagatggatg aggcctttga caacaaaaac      780
55 attccagcgc tcgtctctgc ccttcagaac ggccagactg ttgtctcccg cgaccggttc      840
56 aaggccgtaa cttttagtag aaaccacgac accgatataa tctggaacaa gtatccagcc      900
57 tacgcgttca tcctcaccta cgagggccag ccgacaatat tctaccgcga ctacgaggag      960

```

RAW SEQUENCE LISTING

DATE: 04/01/2003

PATENT APPLICATION: US/10/081,739A

TIME: 13:17:23

Input Set : A:\09010-107001.txt

Output Set: N:\CRF4\04012003\J081739A.raw

```

58 tggctcaaca aggataagct caagaacctc atctggatac atgagaacct cgccggagga 1020
59 agcaccgaca tagtctacta cgataacgat gaactcatct tcgtcaggaa cggctacggg 1080
60 gacaagccgg ggcttataac ctacatcaac ctaggctcga gcaaggccgg aaggtgggtt 1140
61 tatgtgccga agttcgcggg cgcggtgcac cagcagtata ctggtaacct cggaggctgg 1200
62 gtagacaagt acgtctactc aagcggctgg gtctatctcg aagctccagc ttacgacct 1260
63 gccaacgggc agtatggcta ctccgtgtgg agctactgcg ggggtgggctg a 1311

```

65 <210> SEQ ID NO: 2

66 <211> LENGTH: 436

67 <212> TYPE: PRT

68 <213> ORGANISM: Artificial Sequence

70 <220> FEATURE:

71 <223> OTHER INFORMATION: Synthetically engineered

73 <400> SEQUENCE: 2

```

74 Met Ala Lys Tyr Ser Glu Leu Glu Lys Gly Gly Val Ile Met Gln Ala
75 1 5 10 15
76 Phe Tyr Trp Asp Val Pro Ser Gly Gly Ile Trp Trp Asp Thr Ile Arg
77 20 25 30
78 Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp Ile
79 35 40 45
80 Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp
81 50 55 60
82 Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val
83 65 70 75 80
84 Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn Met Ile Asn Thr
85 85 90 95
86 Ala His Ala Tyr Gly Met Lys Val Ile Ala Asp Ile Val Ile Asn His
87 100 105 110
88 Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr
89 115 120 125
90 Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
91 130 135 140
92 Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe
93 145 150 155 160
94 Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp
95 165 170 175
96 Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
97 180 185 190
98 Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp Val
99 195 200 205
100 Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
101 210 215 220
102 Trp Asp Thr Asn Val Asp Ala Val Leu Asn Trp Ala Tyr Ser Ser Gly
103 225 230 235 240
104 Ala Lys Val Phe Asp Phe Ala Leu Tyr Tyr Lys Met Asp Glu Ala Phe
105 245 250 255
106 Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln
107 260 265 270
108 Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
109 275 280 285

```

RAW SEQUENCE LISTING

DATE: 04/01/2003

PATENT APPLICATION: US/10/081,739A

TIME: 13:17:23

Input Set : A:\09010-107001.txt

Output Set: N:\CRF4\04012003\J081739A.raw

```

110 His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
111      290      295      300
112 Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu
113 305      310      315      320
114 Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Glu Asn
115      325      330      335
116 Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu
117      340      345      350
118 Ile Phe Val Arg Asn Gly Tyr Gly Asp Lys Pro Gly Leu Ile Thr Tyr
119      355      360      365
120 Ile Asn Leu Gly Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys
121      370      375      380
122 Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
123 385      390      395      400
124 Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
125      405      410      415
126 Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr
127      420      425      430
128 Cys Gly Val Gly
129      435

```

131 <210> SEQ ID NO: 3

132 <211> LENGTH: 1419

133 <212> TYPE: DNA

134 <213> ORGANISM: Unknown

136 <220> FEATURE:

137 <223> OTHER INFORMATION: Obtained from an environmental sample

139 <400> SEQUENCE: 3

```

140 atgttctctgc tcgctgtttt gctcactgcc tcgctgttct gcccaacagg acagcccgcc      60
141 aaggctgccg caccgtttta cggcaccatg atgcagtatt ttgaatggta cttgccggat      120
142 gatggcacgt tatggaccaa agtggccaat gaagccaaca acttatccag ccttggcatc      180
143 accgctcttt ggctgccgcc cgcttacaaa ggaacaagcc gcagcgacgt agggtagcga      240
144 gtatacgact tgtatgacct cggcgcaattc aatcaaaaag ggaccgtccg cacaaaatac      300
145 ggaacaaaag ctcaatatct tcaagccatt caagccgccc acgccgctgg aatgcaagtg      360
146 tacgccgatg tcgtgttcga ccataaaggc ggcgctgacg gcacggaatg ggtggacgcc      420
147 gtcgaagtca atccgtccga ccgcaaccaa gaaatctcgg gcacctatca aatccaagca      480
148 tggacgaaat ttgattttcc cgggcggggc aacacctact ccagctttta gtggcgctgg      540
149 taccattttg acggcggtga ttgggacgaa agccgaaaat tgagccgcat ttacaaattc      600
150 cgcgcatcg gcaaagcgtg ggattgggaa gtagacacgg aaaacggaaa ctatgactac      660
151 ttaatgtatg ccgaccttga tatggatcat cccgaagtcg tgaccgagct gaaaaactgg      720
152 gggaaatggt atgtcaacac aacgaacatt gatgggttcc ggcttgatgc cgtcaagcat      780
153 attaatgtca gtttttttcc tgattggttg tcgtatgtgc gttctcagac tggcaagccg      840
154 ctattttaccg tcggggaata ttggagctat gacatcaaca agttgcacaa ttacattacg      900
155 aaaacagacg gaacgatgtc tttgtttgat gccccgttac acaacaaatt ttataccgct      960
156 tccaaatcag ggggcgcatt tgatatgcgc acgttaatga ccaatactct catgaaagat      1020
157 caaccgacat tggccgtcac cttcgttgat aatcatgaca ccgaaccggg ccaagcgctg      1080
158 cagtcatggg tcgacccatg gttcaaaccg ttggcttacg cttttattct aactcggcag      1140
159 gaaggatacc cgtgcgtctt ttatggtgac tattatggca ttccacaata taacattcct      1200
160 tcgctgaaaa gcaaaatcga tccgctcctc atcgcgcgca gggattatgc ttacggaacg      1260
161 caacatgatt atcttgatca ctccgacatc atcgggtgga caagggaagg ggtcactgaa      1320

```

RAW SEQUENCE LISTING

DATE: 04/01/2003

PATENT APPLICATION: US/10/081,739A

TIME: 13:17:23

Input Set : A:\09010-107001.txt

Output Set: N:\CRF4\04012003\J081739A.raw

```

162 aaaccaggat cggggtggc cgcactgatc accgatgggc cgggaggaag caaatggatg 1380
163 tactgttggc aaacaacacg ctggaaaagt gttctatga 1419
165 <210> SEQ ID NO: 4
166 <211> LENGTH: 1539
167 <212> TYPE: DNA
168 <213> ORGANISM: Unknown
170 <220> FEATURE:
171 <223> OTHER INFORMATION: Obtained from an environmental sample
173 <400> SEQUENCE: 4
174 atgaaacaac aaaaacggct ttacgcccga ttgctgacgc tgttatttgc gctcatcttc 60
175 ttgctgcctc attctgcagc agcggcggca aatcttaatg ggacgctgat gcagtatttt 120
176 gaatgttaca tgcccaatga cggccaacat tggaagcgct tgcaaaacga ctcgcatat 180
177 ttggctgaac acggtattac tgccgtctgg attcccccg catataaggg aacgagccaa 240
178 gcggatgtgg gctacggtgc ttacgacctt tatgatttag gggagtttca tcaaaaaggg 300
179 acggttcgga caaagtacgg caaaaagga gagctgcaat ctgcatcaa aagtcttcat 360
180 tcccgcgaca ttaacgttta cggggatgtg gtcacacacc acaaaggcgg cgctgatgcg 420
181 accgaagatg taaccgcggg tgaagtcatg cccgctgacc gcaaccgcgt aatttcagga 480
182 gaacaccgaa ttaaagcctg gacacatttt cattttccgg gccgcggcag cacatacagc 540
183 gatttttaaat ggcattggta ccattttgac ggaaccgatt gggacgagtc ccgaaagctg 600
184 aaccgcatct ataagtttca aggaaaggct tgggattggg aagtttcaa tgaaaacggc 660
185 aactatgatt atttgatgta tgccgacatc gattatgacc atcctgatgt cgcagcagaa 720
186 attaagagat ggggcacttg gtatgccaat gaactgcaat tggacggttt ccgtcttgat 780
187 gctgtcaaac acattaaatt ttcttttttg cgggattggg ttaatcatgt cagggaaaaa 840
188 acggggaagg aaatgtttac ggtagctgaa tattggcaga atgacttggg cgcgctggaa 900
189 aactatttga acaaaacaaa ttttaatcat tcagtgtttg acgtgccgct tcattatcag 960
190 ttccatgctg catcgacaca gggaggcggc tatgatatga ggaaattgct gaacggtacg 1020
191 gtcgtttcca agcatccgtt gaaagcgggt acatttgcg ataaccatga tacacagccg 1080
192 gggcaatcgc ttgagtcgac tgtccaaaca tggtttaagc cgcttgctta cgctttcatt 1140
193 ctcaaaaggg aatctggata ccctcagggt ttctacgggg atatgtacgg gacgaaagga 1200
194 gactcccagc gcgaaattcc tgccttgaaa caaaaattg aaccgatctt aaaagcgaga 1260
195 aaacagtatg cgtacggagc acagcatgat tatttcgacc accatgacat tgtcggctgg 1320
196 acaagggaag gcgacagctc ggttgcaaat tcaggtttgg cggcattaat aacagacgga 1380
197 cccggtgggg caaagcgaat gtatgtcggc cggcaaaacg ccggtgagac atggcatgac 1440
198 attaccgga accgttcgga gccggtgtgc atcaattcgg aaggctgggg agagtttcac 1500
199 gtaaacggcg gtcggtttc aatttatgtt caaagatag 1539
201 <210> SEQ ID NO: 5
202 <211> LENGTH: 1395
203 <212> TYPE: DNA
204 <213> ORGANISM: Unknown
206 <220> FEATURE:
207 <223> OTHER INFORMATION: Obtained from an environmental sample
209 <400> SEQUENCE: 5
210 gtggtgcaca tgaagttgaa gtaccttgcc ttagttttgt tggctgtggc ttcgataggc 60
211 ctactctcga ctccagtggg tgctgccaa g tactccgaac tcgaagaggg cggtgttata 120
212 atgcaggcct tctactggga tgttcccga gggggaatct ggtgggacac cataagacag 180
213 aaaatcccgg agtggtacga cgctggaatc tcggcgatat ggattcctcc agctagcaaa 240
214 gggatgggcg gtggttattc catgggttac gatccctacg atttctttga cctcggcgag 300
215 tactatcaga agggaacagt tgagacgcgc ttcggctcaa aggaggaact ggtgaacatg 360
216 ataaacaccg cacactccta tggcataaag gtgatagcgg acatagtcac aaaccaccgc 420

```

RAW SEQUENCE LISTING

DATE: 04/01/2003

PATENT APPLICATION: US/10/081,739A

TIME: 13:17:23

Input Set : A:\09010-107001.txt

Output Set: N:\CRF4\04012003\J081739A.raw

```

217 gccggtggag accttgagtg gaaccctttt gtaaacaact atacttggac agacttctcc 480
218 aaggtcgctt ccggtaaata cacggccaac taccttgact tccacccaaa cgaggtcaag 540
219 tgctgcatg aggttacatt tggtagcttt ccgacatcg ccacagagaa gagctgggat 600
220 cagtactggc tctgggcaag caatgagagc tacgccgat atctccggag catagggatc 660
221 gatgcatggc gtttcgacta cgtcaaaggc tacggagcgt gggttgttaa tgactggctc 720
222 agctggtggg gaggtctggc cgttgagag tactgggaca cgaacgttga tgcactcctt 780
223 aactgggcat acgacagcgg tgccaaggct tttgacttcc cgctctacta caagatggac 840
224 gaagcctttg acaacaccaa catccccgct ttggtttacg ccctccagaa cggaggaaca 900
225 gtcgtttccc gcgatccctt caaggcagta actttcgttg ccaaccacga tacagatata 960
226 atctggaaca agtatccggc ttatgcgttc atccttacct atgagggaca gcctgttata 1020
227 ttttaccgag actacgagga gtggctcaac aaggataagc ttaacaacct tatctggata 1080
228 cacgagcacc ttgccggagg aagtaccaag atcctctact acgataacga tgagctaata 1140
229 ttcagtaggg agggctacgg gagcaagccg ggcctcataa cctacataaa cctcggaaac 1200
230 gactgggccg agcgtggttg gaacgtcggc tcaaagtttg ccggctacac aatccatgaa 1260
231 tacacaggca atctcgggtg ctgggttgac aggtgggttc agtacgatgg atgggttaaa 1320
232 ctgacggcac ctctcatga tccagccaac ggatattacg gctactcagt ctggagctac 1380
233 gcaggcgtcg gatga 1395
235 <210> SEQ ID NO: 6
236 <211> LENGTH: 1386
237 <212> TYPE: DNA
238 <213> ORGANISM: Unknown
240 <220> FEATURE:
241 <223> OTHER INFORMATION: Obtained from an environmental sample
243 <400> SEQUENCE: 6
244 atgaagaagt ttgtcgccct gttcataacc atgttttttcg tagtgagcat ggcagtcggt 60
245 gcacagccag ctacgcccgc aaagtattcc gagctcgaag aaggcggcgt tataatgcag 120
246 gccttctact gggacgtccc aggtggagga atctggtggg acaccatcag gagcaagata 180
247 ccggagtggg acgaggcggg aatatccgcc atttggtatc cgccagccag caaggggatg 240
248 agcggcgggt actcgatggg ctacgatccc tacgatttct ttgacctcgg cgagtacaac 300
249 cagaagggaa ccatcgaaac gcgctttggc tctaaacagg agctcatcaa tatgataaac 360
250 acggccccatg cctacggcat aaaggtcata gcggacatcg tcataaacca ccgcgcaggc 420
251 ggagacctcg agtggaaacc gttcgttggg gactacacct ggacggactt ctcaaagggtg 480
252 gcctcgggca aatatactgc caactacctc gacttccacc ccaacgaggt caagtgtgtg 540
253 gacgagggca catttgaggg cttcccagac atagcccacg agaagagctg ggaccagcac 600
254 tggctctggg cgagcgtatg gagctacgcc gcctacctaa ggagcatcgg cgttgatgcc 660
255 tggcgctttg actacgtgaa gggctacgga gcgtgggtcg tcaaggactg gctcaactgg 720
256 tggggcggct gggccgttgg cgagtactgg gacaccaacg ttgatgcact cctcaactgg 780
257 gcctactcga gcggcgccaa ggtcttcgac ttcccgctct actacaagat ggatgaggcc 840
258 tttgacaaca aaaacattcc agcgtcgtc tctgcccttc agaacggcca gactgttgct 900
259 tccgcgacc cgttcaaggc cgtaaccttt gtagcaaacc acgacaccga tataatctgg 960
260 aacaagtacc ttgcttatgc tttcatcctc acctacgaag gccagcccgt catattctac 1020
261 cgcgactacg aggagtggct caacaaggac aggttgaaca acctcatatg gatacacgac 1080
262 cacctcgag gtggaagcac gagcatagtc tactacgaca gcgacgagat gatcttcgtg 1140
263 aggaacggct atggaagcaa gcctggcctt ataacttaca tcaacctcgg ctcgagcaag 1200
264 gttggaagg ggttttatgt gccgaagttc gcgggcgcgt gcatccacga gtatactgg 1260
265 aacctcggag gctgggtaga caagtacgtc tactcaagcg gctgggtcta tctcgaagct 1320
266 ccagcttacg accctgccaa cgggcagtat ggctactccg tgtggagcta ttgcggtgtt 1380
267 ggtga 1386
269 <210> SEQ ID NO: 7

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/081,739A

DATE: 04/01/2003

TIME: 13:17:24

Input Set : A:\09010-107001.txt

Output Set: N:\CRF4\04012003\J081739A.raw